

CONVERGENCE OF MOMENTS OF STANDARDIZED QUANTILES

BY

KEAVEN ANDERSON

TECHNICAL REPORT NO. 4

DECEMBER 15, 1980

U.S. ARMY RESEARCH OFFICE
RESEARCH TRIANGLE PARK, NORTH CAROLINA
CONTRACT NO. DAAG29-79-C-0166

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA



CONVERGENCE OF MOMENTS OF STANDARDIZED QUANTILES

BY

KEAVEN ANDERSON

TECHNICAL REPORT NO. 4

DECEMBER 15, 1980

U.S. ARMY RESEARCH OFFICE
RESEARCH TRIANGLE PARK, NORTH CAROLINA
CONTRACT NO. DAAG29-79-C-0166

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

CONVERGENCE OF MOMENTS OF STANDARDIZED QUANTILES

Keaven Anderson

Abstract. A sequence of independent identically distributed random variables is considered. Necessary and sufficient conditions on the density of the distribution are given for convergence of the moments of standardized quantiles of the first n observations as $n \rightarrow \infty$. Similar conditions are given for the convergence of the moment generating function.

Running title: Convergence of moments of quantiles.

AMS 1980 subject classification numbers: Primary 60 G30; Secondary 60 F25.

Keywords and phrases: order statistics; quantiles; moments; moment generating functions; asymptotics; L^p -convergence.

§1. Introduction and summary. Let $F(x)$, $-\infty < x < \infty$, be a distribution function with corresponding density $f(x)$. Assume $c \in (0, 1)$, $F(q) = c$ and that the derivative of $F(x)$ exists and is positive at $x = q$. Let X_1, X_2, \dots be independent, identically distributed random variables with distribution function $F(x)$. Denote the order statistics of X_1, X_2, \dots, X_n by $X_{1:n}, X_{2:n}, \dots, X_{n:n}$. Let $a_n = cn + O(1)$. Assume $Z \sim N(0, c(1-c)/f^2(q))$. Wretman (1978) has shown that $n^{1/2}(X_{a_n:n} - q)$ converges in distribution to Z . Other authors have shown the same result under additional assumptions on $f(x)$; e. g. Rao (1973) assumed $f(x)$ continuous at q and Cramér (1946) assumed the derivative of $f(x)$ continuous at q .

This paper gives necessary and sufficient conditions for $E\{g(n^{1/2}(X_{a_n:n} - q))\}$ to converge to $E\{g(Z)\}$ for a class of functions which includes $g(x) = x^k$ and $g(x) = e^{tx}$. One pair of necessary and sufficient conditions for $g(x)$ in this class is

$$\liminf_{x \rightarrow \infty} \frac{-\log(1 - F(x))}{\log(|g(x)|)} > 0 \text{ and } \liminf_{x \rightarrow \infty} \frac{-\log(F(-x))}{\log(|g(-x)|)} > 0. \quad (1)$$

Another necessary and sufficient condition is that there exists $\delta > 0$ such that $E\{|g(X_1)|^\delta\} < \infty$.

A theorem which will eventually connect these two types of conditions will be given in section 2. This theorem is actually of some interest in itself as it may be used to determine whether or not the expectation of a function of a random variable exists. The results stated above will be proved in section 3.

Besides the obvious applications, applications to sequential occupancy and related problems are suggested by Holst (1981) and Anderson *et. al.* (1980).

§2. Tails and expectations. A "nearly" necessary and sufficient condition for the mean of a distribution to exist will be given in this section. This result is easily extended to show whether or not the expectations of many functions of a random variable exist. As examples of extensions moments and the moment generating function will be considered. Lacking a suitable reference a complete proof of theorem 1 is given here.

Theorem 1. Suppose X is an arbitrary non-negative random variable with distribution function $F(x)$. Let

$$\alpha = \liminf_{x \rightarrow \infty} \frac{-\log(1 - F(x))}{\log x}.$$

If $\alpha < 1$ then $E\{X\} = \infty$. If $\alpha > 1$ then $E\{X\} < \infty$. If $\alpha = 1$ it may be the case that $E\{X\} = \infty$ or $E\{X\} < \infty$.

Proof. The last part of the proposition will be proved first. If $F(x) = 1 - 1/x$ for $x \geq 1$ then $\alpha = 1$ and $E\{X\} = \infty$. If $F(x) = 1 - x^{-1}e^{-\sqrt{\log x}}$ for $x \geq 1$ then $\alpha = 1$ and $E\{X\} = 2$.

Assume $\alpha > 1$. Then there exists $\lambda \in (1, \alpha)$ and x_1 such that for $x \geq x_1$ $-\log(1 - F(x)) > \lambda \log x$. This implies that for $x \geq x_1$ $1 - F(x) < x^{-\lambda}$ and thus $E\{X\} < \infty$.

Now suppose $\alpha < 1$. It will be shown that $\sum_{k=1}^{\infty} P\{X > k\} = \infty$ which implies $E\{X\} = \infty$. If $\alpha < 1$ then there exists $\lambda \in (\alpha, 1)$ and $x_1 < x_2 < \dots$ such that $1 - F(x_n) > x_n^{-\lambda}$, $n = 1, 2, \dots$ and $x_n \rightarrow \infty$ as $n \rightarrow \infty$. Let y_n be the greatest integer less than or equal to x_n , $n = 1, 2, \dots$ and let $y_0 = 0$. Then

$$\sum_{k=1}^{\infty} P\{X > k\} = \sum_{n=1}^{\infty} \sum_{k=y_{n-1}+1}^{y_n} P\{X > k\} \geq \sum_{n=1}^{\infty} \frac{y_n - y_{n-1}}{x_n^{\lambda}} \geq \lim_{x \rightarrow \infty} \frac{y_n}{x_n^{\lambda}} = \infty.$$

Definition. For an arbitrary non-decreasing function $g(x)$ define $g^{-1}(x) = \inf\{y : g(y) \geq x\}$.

Corollary 1.1. Let X be an arbitrary random variable and denote its distribution function by $F(x)$. Let $g(x)$ be an arbitrary non-decreasing, non-negative function such that $g(x) \rightarrow \infty$ as $x \rightarrow F^{-1}(1)$. Let

$$\alpha = \liminf_{x \rightarrow F^{-1}(1)} \frac{-\log(1 - F(x))}{\log g(x)}.$$

If $\alpha > 1$ then $E\{g(X)\} < \infty$, if $\alpha < 1$ then $E\{g(X)\} = \infty$, and if $\alpha = 1$ it may be the case that $E\{g(X)\} = \infty$ or $E\{g(X)\} < \infty$.

Proof. This follows from theorem 1 since

$$\begin{aligned}\liminf_{x \rightarrow F^{-1}(1)} \frac{-\log(1 - F(x))}{\log g(x)} &= \liminf_{x \rightarrow \infty} \frac{-\log(1 - F(g^{-1}(x)))}{\log x} \\ &= \liminf_{x \rightarrow \infty} \frac{-\log P\{g(X) > x\}}{\log x}.\end{aligned}$$

Corollary 1.2. For an arbitrary random variable X with distribution function $F(x)$ let

$$\alpha_+ = \liminf_{x \rightarrow \infty} \frac{-\log(1 - F(x))}{\log x}, \quad \alpha_- = \liminf_{x \rightarrow \infty} \frac{-\log F(-x)}{\log x},$$

and $\alpha = \min(\alpha_+, \alpha_-)$. If $0 \leq k < \alpha$ then $E\{|X|^k\} < \infty$. If $k > \alpha$ then $E\{|X|^k\} = \infty$. For $k = \alpha$ it may be the case that either $E\{|X|^k\} = \infty$ or $E\{|X|^k\} < \infty$.

Corollary 1.3. For an arbitrary random variable X with distribution function $F(x)$ let

$$\alpha_2 = \liminf_{x \rightarrow \infty} \frac{-\log(1 - F(x))}{x} \quad \text{and} \quad \alpha_1 = -\liminf_{x \rightarrow \infty} \frac{-\log F(-x)}{x}.$$

If $t \in (\alpha_1, \alpha_2)$ then $E\{e^{tX}\} < \infty$. If $t < \alpha_1$ or $t > \alpha_2$ then $E\{e^{tX}\} = \infty$.

§3. Moments of order statistics. We are now prepared to address the questions of interest. The proof of the necessity of (1) for the existence of moments of standardized quantiles is now trivial in many cases. The proofs of the convergence results presented here are analytic and somewhat tedious in nature. Following is the most general result concerning "necessity" that will be given.

Theorem 2. Suppose $g(x)$ is a non-decreasing, non-negative function, X_1, X_2, \dots are independent identically distributed random variables with distribution function $F(x)$, and that

$$\liminf_{x \rightarrow \infty} \frac{-\log(1 - F(x))}{\log g(x)} = 0.$$

Then for any $\beta > 1$, α and non-negative integers $n, k, 1 \leq k \leq n$

$$E\{g(\beta(X_{k:n} - \alpha))\} = \infty.$$

Proof. This follows from theorem 1 since for large x

$$\begin{aligned} P\{g(\beta(X_{k:n} - \alpha)) > x\} &\geq P\{g(X_{k:n}) > x\} \\ &\geq \binom{n}{k-1} (1 - F(g^{-1}(x))^{n-k+1} (F(g^{-1}(x)))^{k-1} \end{aligned}$$

and thus

$$\begin{aligned} \liminf_{x \rightarrow \infty} \frac{-\log P\{g(\beta(X_{k:n} - \alpha)) > x\}}{\log x} &\leq (n - k + 1) \liminf_{x \rightarrow \infty} \frac{-\log(1 - F(g^{-1}(x)))}{\log x} \\ &\leq (n - k + 1) \liminf_{x \rightarrow \infty} \frac{-\log(1 - F(x))}{\log g(x)}. \end{aligned}$$

For the remainder of the paper the assumptions of the first paragraph of section 1 will be used. Convergence results will be shown using the integral representation

$$\begin{aligned} E\{g(n^{1/2}(X_{a_n:n} - q))\} \\ = \int_{-\infty}^{\infty} g(n^{1/2}(x - q)) n \binom{n-1}{a_n-1} F(x)^{a_n-1} (1 - F(x))^{n-a_n} f(x) dx. \end{aligned} \quad (2)$$

Letting $c_n = a_n/n$ it is easy to show using Stirling's formula that

$$\binom{n-1}{a_n-1} \sim (2\pi n)^{-1/2} c_n^{-nc_n+1/2} (1 - c_n)^{-n(1-c_n)-1/2}. \quad (3)$$

Letting

$$p_n(u) = c_n(\log c_n - \log u) + (1 - c_n)(\log(1 - c_n) - \log(1 - u))$$

$n = 1, 2, \dots, \infty$, where $c_\infty = c$, it follows that as $n \rightarrow \infty$

$$\begin{aligned} E\{g(n^{1/2}(X_{a_n:n} - q))\} \\ \sim \int_{-\infty}^{\infty} \left(\frac{nc_n}{2\pi(1 - c_n)} \right)^{1/2} \frac{f(x)}{F(x)} \exp\left(\log g(n^{1/2}(x - q)) - np_n(F(x))\right) dx. \end{aligned} \quad (4)$$

Three assumptions will be made in much of the following:

- i) $g(x)$ is a continuous, non-decreasing, non-negative function,
- ii) there exists $\beta > 0$ such that if $t > 1, x > 0$ $\log g(tx) < t\beta \log g(x)$,
- iii) $\liminf_{x \rightarrow \infty} \frac{-\log(1 - F(x))}{\log g(x)} > 0$.

Lemma 1. Under assumptions i) – iii) if $\epsilon > 0$ and $\gamma < 1/4$

$$0 = \lim_{n \rightarrow \infty} \int_{q + \epsilon n^{-\gamma}}^{\infty} n^{1/2} \frac{f(x)}{F(x)} \exp\left(\log g(n^{1/2}(x - q)) - np_n(F(x))\right) dx.$$

Proof. Without loss of generality we assume $\gamma > 0$. Since

$$p'_n(u) = \frac{-c_n}{u} + \frac{1 - c_n}{1 - u} \quad \text{and} \quad p''_n(u) = \frac{c_n}{u^2} + \frac{1 - c_n}{(1 - u)^2}$$

it follows that $p_n(u)$ is convex and assumes its minimum of 0 at c_n . This implies that for any $u > c_n$ $p_n(u) \rightarrow p_\infty(u) > 0$. Since for a given $x > q$ and n large $\log(g(n^{1/2}(x - q))) < \beta n^{1/2} \log(g(x - q))$ it follows that the integrand of the lemma goes pointwise to zero.

The integrand will now be bounded in the tail for large n . Let $\eta > 0$ and $x_1 > q$ be such that for if $x > x_1$ then $-\log(1 - F(x)) > \eta \log g(x)$. It follows now that there exists k_1 such that for $x > x_1$ the integrand is bounded by

$$k_1 f(x) \exp\left(n^{1/2} \beta \log g(x - q) + \log n^{1/2} - n \eta \log g(x)\right).$$

This can be bounded by $k'_1 f(x)$ for some $k'_1 > 0$.

The integrand on some interval of the form $[q + \epsilon/n^\gamma, x_1]$ will now be considered. Using the first order Taylor's series expansion with remainder it follows that for $u \geq c_n$, some $\theta_n(u)$ such that $\theta_n(u) \in [c_n, u]$

$$p_n(u) = \left(\frac{c_n}{\theta_n^2(u)} + \frac{1 - c_n}{(1 - \theta_n(u))^2} \right) \frac{(u - c_n)^2}{2} \geq c_n \frac{(u - c_n)^2}{2}.$$

For n large $F(x) > c + (x - q)f(q)/2$ for $x \in (q, q + \epsilon/n^\gamma]$. Thus the integrand is bounded for some $k_2 > 0$ by

$$k_2 f(x) \exp\left(\log n^{1/2} + \log g(n^{1/2}(x_1 - q)) - nc_n \epsilon^2 f^2(q)/(4n^{2\gamma})\right).$$

Since $\gamma < 1/4$ and $\log g(n^{1/2}(x_1 - q)) = O(n^{1/2})$ it follows that for some k'_2 and N the integrand is bounded by $k'_2 f(x)$ for $n \geq N$.

The contention now follows by the dominated convergence theorem since the integrand is bounded by $\max(k'_1, k'_2) f(x)$ for n sufficiently large.

Lemma 2. Assume $Y \sim N(0, c(1 - c))$. For any $\epsilon > 0$ as $n \rightarrow \infty$

$$\xi_n(\epsilon) = \int_{c - \epsilon/n^{1/2}}^{c + \epsilon/n^{1/2}} g(n^{1/2}(u - c)) n \binom{n-1}{a_n-1} u^{a_n-1} (1-u)^{n-a_n} du \rightarrow E\{g(Y)\}.$$

Proof. Letting $z = n^{1/2}(u - c_n)$ and applying equation (3)

$$\xi_n(\epsilon) \sim \int_{c - \epsilon n^{3/10}}^{c + \epsilon n^{3/10}} \frac{g(z + n^{1/2}(c_n - c))}{(2\pi c_n(1 - c_n))^{1/2}} \left(1 + \frac{z}{n^{1/2}c_n}\right)^{a_n - 1} \left(1 - \frac{z}{n^{1/2}(1 - c_n)}\right)^{n - a_n} dz.$$

Using the Taylor's series expansion

$$\log(1 + z) = z - \frac{z^2}{2} + \frac{z^3}{3(1 + \theta(x))^3}$$

which is valid for $-1 < z < 1$ and some $\theta(x)$ between 0 and x . it is a straightforward calculation to show that

$$\max_{|z - c| < \epsilon n^{3/10}} \left| \log \left(\left(1 + \frac{z}{n^{1/2}c_n}\right)^{a_n - 1} \left(1 - \frac{z}{n^{1/2}(1 - c_n)}\right)^{n - a_n} \right) + \frac{z^2}{2c(1 - c)} \right| \rightarrow 0$$

as $n \rightarrow \infty$. Thus

$$\xi_n(\epsilon) \sim \int_{c - \epsilon n^{3/10}}^{c + \epsilon n^{3/10}} \frac{g(z + n^{1/2}(c_n - c))}{(2\pi c(1 - c))^{1/2}} \exp(-z^2/(2c(1 - c))) dz.$$

Since $c_n = c + O(1/n)$ assumption i) implies $g(z + n^{1/2}(c_n - c)) \rightarrow g(z)$. Assumption ii) implies that $g(z)$ grows at most exponentially. This implies that the integrand above can be bounded by $g(z)k_1 \exp(-k_2 z^2)$ for some k_1, k_2 and n large. Thus by the dominated convergence theorem $\xi(\epsilon) \rightarrow E\{g(Y)\}$.

We are now prepared to prove the most general convergence result which will be given here.

Theorem 3. Assuming i)-iii), $f(q) > 0$, and that $Z \sim N(0, c(1 - c)/f^2(q))$

$$E\{g(n^{1/2}(X_{a_n:n} - q))\} \rightarrow E\{g(Z)\}.$$

Proof. From (2)-(4), and lemma 1 and its obvious analog it follows that for any $\epsilon > 0$

$$\begin{aligned} & E\{g(n^{1/2}(X_{a_n:n} - q))\} \\ & \sim \int_{F^{-1}(c - \epsilon)}^{F^{-1}(c + \epsilon)} g(n^{1/2}(x - q)) n \binom{n-1}{a_n-1} F(x)^{a_n-1} (1 - F(x))^{n-a_n} f(x) dx. \end{aligned}$$

Letting $u = F(x)$ the above integral is equal to

$$\int_{c-\epsilon}^{c+\epsilon} g(n^{1/2}(F^{-1}(u) - q)) n \binom{n-1}{a_n-1} u^{a_n-1} (1-u)^{n-a_n} du.$$

For any $\delta > 0$ there exists $\epsilon > 0$ such that for $u \in (c - \epsilon, c + \epsilon)$

$$(1 - \delta) \frac{u - c}{f(q)} < F^{-1}(u) - q < (1 + \delta) \frac{u - c}{f(q)}.$$

The result now follows from lemmas 1 and 2 after substituting the two bounds into the integral above.

Propositions 1 and 3 below follow from theorems 2 and 3. Proposition 2 follows from proposition 1 and corollary 1.2. Proposition 4 follows from proposition 3 and corollary 1.3. For each of the following propositions recall that we assume the assumptions of the first paragraph of section 1.

Proposition 1. Define α_+ and α_- as in corollary 1.2. If $k > 0$ then $E\{(n^{1/2}(X_{a_n:n} - q))^k\} \rightarrow E\{Z^k\}$ as $n \rightarrow \infty$ if and only if $\alpha_+ > 0$ and $\alpha_- > 0$.

Proposition 2. If $k > 0$ then $E\{(n^{1/2}(X_{a_n:n} - q))^k\} \rightarrow E\{Z^k\}$ as $n \rightarrow \infty$ if and only if there exists $\delta > 0$ such that $E\{|X_1|^\delta\} < \infty$.

Proposition 3. Define α_1 and α_2 as in corollary 1.3. Then for all $t > 0$ $E\{\exp(tn^{1/2}(X_{a_n:n} - q))\} \rightarrow E\{e^{tZ}\}$ if and only if $\alpha_2 > 0$. Similarly for all $t < 0$ $E\{\exp(tn^{1/2}(X_{a_n:n} - q))\} \rightarrow E\{e^{tZ}\}$ if and only if $\alpha_1 < 0$.

Proposition 4. For all $t > 0$ $E\{\exp(tn^{1/2}(X_{a_n:n} - q))\} \rightarrow E\{e^{tZ}\}$ if and only if there exists $\epsilon > 0$ such that $E\{e^{\epsilon X_1}\} < \infty$. Similarly for all $t < 0$ $E\{\exp(tn^{1/2}(X_{a_n:n} - q))\} \rightarrow E\{e^{tZ}\}$ if and only if there exists $\epsilon > 0$ such that $E\{e^{-\epsilon X_1}\} < \infty$.

References

Anderson, K., Sobel, M. and Uppuluri, V. R. R., (1980), "Quota fulfillment times," *Stanford University Department of Statistics Technical Report No. xxx*.

Cramér, H., (1946), *Mathematical Methods of Statistics*, Princeton University Press.

Holst, Lars, (1981), "On sequential occupancy problems," *J. Appl. Prob.*, to appear June, 1981.

Rao, C.R. (1973), *Linear Statistical Inference and Its Applications*, Wiley, New York.

Wretman, J., (1978) "A simple derivation of the asymptotic distribution of a sample quantile," *Scandinavian Journal of Statistics* 5 123-124.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER No. 4	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) CONVERGENCE OF MOMENTS OF STANDARDIZED QUANTILES		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Keaven Anderson		8. CONTRACT OR GRANT NUMBER(s) ARO DAAG29-79-C-0166
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, California		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P. O. Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE December 15, 1980
		13. NUMBER OF PAGES 7
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES The view, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Order statistics; quantiles; moments; moment generating functions; asymptotics; L^p -convergence.		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) A sequence of independent identically distributed random variables is considered. Necessary and sufficient conditions on the density of the distribution are given for convergence of the moments of standardized quantiles of the first n observations as $n \rightarrow \infty$. Similar conditions are given for the convergence of the moment generating function.		

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)